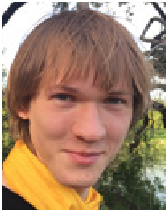
Statistical properties of random lambda-terms in de-Bruijn notation*

Maciej Bendkowski[3] Olivier Bodini[1] Sergey Dovgal[1,2,4]

[1]Université Paris-13, [2]Université Paris-Diderot [3]Jagiellonian University [4]Moscow Institute of Physics and Technology

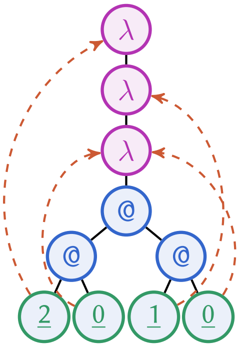


CLA-2017, Götenburg, Sweden

∗ in progress

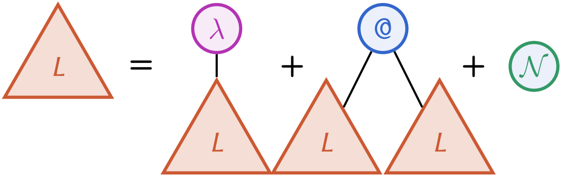
Outline

Example of lambda-term in de-Bruijn notation



A closed lambda-term $\lambda x.\lambda y.\lambda z.xz(yz)$

Grammar of plain lambda-terms



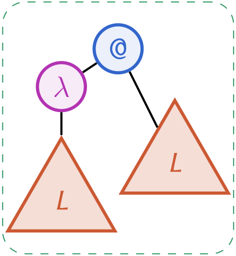
L — plain lambda-term

λ — abstraction

@ — application

$\mathcal{N}$ — variable; de Bruijn index $\in \{0, 1, 2, \ldots\}$

Redex and beta-reduction

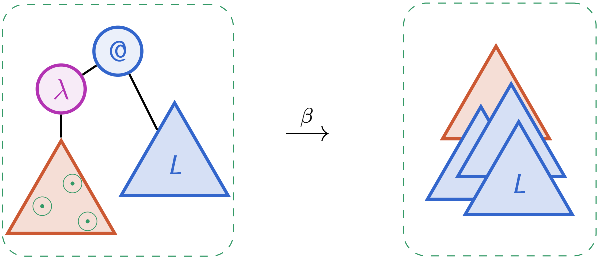


$$(\lambda n.n \times 2)7 \xrightarrow{\beta} 7 \times 2$$

$$\Omega = (\lambda x.xx)(\lambda x.xx)$$

$$\Omega \xrightarrow{\beta} \Omega$$

Redex and beta-reduction



Main question

Investigate statistical properties of random plain / closed lambda-terms in de-Bruijn notation:

- number of lambdas, variables, abstractions,. . .
- length to the leftmost outermost *redex*,
- unary height, longest lambda-run
- . . .

Statistical properties

Random generation

$\Rightarrow$ Property-based testing

Size notion of lambda-terms

- [Bodini, Gardy, Gittenberger, Jacquot '13]
  Closed lambda-terms with variable size $= 1$.
- [David, Grygiel, Kozik, Raffalli, Theyssier, Zaionc '13]
  Closed lambda-terms with variable size $= 0$.
- [Gittenberger, Gołębiewski '16]
  Natural counting of lambda-terms.

$$|0| = a, \quad |S| = b, \quad |\lambda| = d, \quad |@| = d$$

Size notion of lambda-terms

- [Bodini, Gardy, Gittenberger, Jacquot '13]
  Closed lambda-terms with variable size $= 1$.
- [David, Grygiel, Kozik, Raffalli, Theyssier, Zaionc '13]
  Closed lambda-terms with variable size $= 0$.
- [Gittenberger, Gołębiewski '16]
  Natural counting of lambda-terms.

$$|0| = a, \quad |S| = b, \quad |\lambda| = d, \quad |@| = d$$

Size notion of lambda-terms

- [Bodini, Gardy, Gittenberger, Jacquot '13]
  Closed lambda-terms with variable size $= 1$.
- [David, Grygiel, Kozik, Raffalli, Theyssier, Zaionc '13]
  Closed lambda-terms with variable size $= 0$.
- [Gittenberger, Gołębiewski '16]
  Natural counting of lambda-terms.

$$|0| = a, \quad |S| = b, \quad |\lambda| = d, \quad |@| = d$$

What is "random"?

Thanks to previous talks

- *Natural size notion* of lambda-term.

  Stay tuned for the definition and comparison to other models.

- Sample lambda-terms of size n **uniformly**

  (*leitmotif* of this talk, but not the only possibility)

- Sample lambda-terms of size n and parameter value k uniformly.

  Bivariate generating function + tuning of Boltzmann sampler.

- Choose *any subset* of parameters, fix their values and sample at uniform from the desired set.

  [Bodini, Ponty '10]: Newton iteration or asymptotic approximations

  [Bodini, D. '17]: Fast exact tuning in $O\Big(\log^2(\text{size}) \cdot \#\text{vars}^7 \cdot (\#\text{vars}+\#\text{eqs})\Big)$.

What is "random"?

Thanks to previous talks

- *Natural size notion* of lambda-term.

  Stay tuned for the definition and comparison to other models.

- Sample lambda-terms of size n **uniformly**

  (*leitmotif* of this talk, but not the only possibility)

- Sample lambda-terms of size n and parameter value k uniformly.

  Bivariate generating function + tuning of Boltzmann sampler.

- Choose *any subset* of parameters, fix their values and sample at uniform from the desired set.

  [Bodini, Ponty '10]: Newton iteration or asymptotic approximations

  [Bodini, D. '17]: Fast exact tuning in $O\Big(\log^2(\text{size}) \cdot \#\text{vars}^7 \cdot (\#\text{vars}{+}\#\text{eqs})\Big)$.

What is “random”?

Thanks to previous talks

- *Natural size notion* of lambda-term.
  Stay tuned for the definition and comparison to other models.
- Sample lambda-terms of size n **uniformly**
  (*leitmotif* of this talk, but not the only possibility)
- Sample lambda-terms of size n and parameter value k uniformly.
  Bivariate generating function + tuning of Boltzmann sampler.
- Choose *any subset* of parameters, fix their values and sample at uniform from the desired set.
  [Bodini, Ponty ’10]: Newton iteration or asymptotic approximations
  [Bodini, D. ’17]: Fast exact tuning in $O\Big(\log^2(\text{size}) \cdot \#\text{vars}^7 \cdot (\#\text{vars}{+}\#\text{eqs})\Big)$.

What is "random"?

Thanks to previous talks

- *Natural size notion* of lambda-term.

  Stay tuned for the definition and comparison to other models.

- Sample lambda-terms of size n **uniformly**

  (*leitmotif* of this talk, but not the only possibility)

- Sample lambda-terms of size n and parameter value k uniformly.

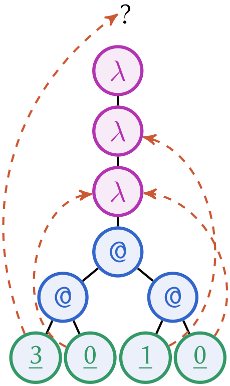
  Bivariate generating function + tuning of Boltzmann sampler.

- Choose *any subset* of parameters, fix their values and sample at uniform from the desired set.

  [Bodini, Ponty '10]: Newton iteration or asymptotic approximations

  [Bodini, D. '17]: Fast exact tuning in $O\left(\log^2(\text{size}) \cdot \#\text{vars}^7 \cdot (\#\text{vars}+\#\text{eqs}) \right)$.

Closed lambda-terms?



The value of each index shouldn't exceed maximal unary distance to parent lambda.

m-open lambda-terms

- **Def.** A lambda-term T is *m-open* if $\lambda^m T$ is closed.
- **Observation.** $m = 0$ corresponds to closed terms
- **Def.** L_m – class of m-open lambda-terms.
- **Def.** L_∞ – class of plain lambda-terms.
- **Observation.** $L_0 \subset L_1 \subset L_2 \subset \ldots \subset L_\infty$

m-open lambda-terms

- **Def.** A lambda-term T is *m-open* if $\lambda^m T$ is closed.
- **Observation.** $m = 0$ corresponds to closed terms
- **Def.** L_m — class of m-open lambda-terms.
- **Def.** L_∞ — class of plain lambda-terms.
- **Observation.** $L_0 \subset L_1 \subset L_2 \subset \ldots \subset L_\infty$

m-open lambda-terms

- **Def.** A lambda-term T is *m-open* if $\lambda^m T$ is closed.
- **Observation.** $m = 0$ corresponds to closed terms
- **Def.** L_m — class of m-open lambda-terms.
- **Def.** L_∞ — class of plain lambda-terms.
- **Observation.** $L_0 \subset L_1 \subset L_2 \subset \ldots \subset L_\infty$

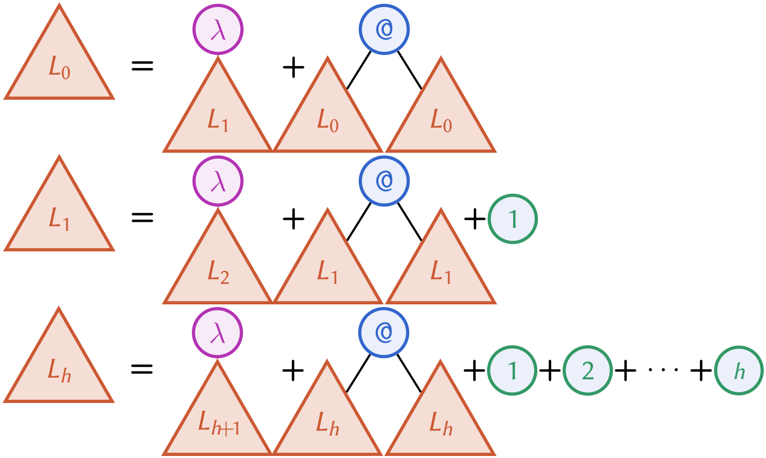
m-open lambda-terms

- **Def.** A lambda-term T is *m-open* if $\lambda^m T$ is closed.
- **Observation.** $m = 0$ corresponds to closed terms
- **Def.** L_m — class of m-open lambda-terms.
- **Def.** L_∞ — class of plain lambda-terms.
- **Observation.** $L_0 \subset L_1 \subset L_2 \subset \ldots \subset L_\infty$

m-open lambda-terms

- **Def.** A lambda-term T is *m-open* if $\lambda^m T$ is closed.
- **Observation.** $m = 0$ corresponds to closed terms
- **Def.** L_m — class of m-open lambda-terms.
- **Def.** L_∞ — class of plain lambda-terms.
- **Observation.** $L_0 \subset L_1 \subset L_2 \subset \ldots \subset L_\infty$

Closed terms specification



Asymptotic number of plain lambda-terms

Theorem. As $n \to \infty$, the number of plain lambda-terms of size n is asymptotically

$$\frac{b_\infty n^{-3/2}}{2\sqrt{\pi}} \left(\frac{1}{\rho}\right)^n, \quad (1-\rho)^3 = 4\rho^2 \ .$$

$$\rho \approx 0.29559$$

Asymptotic number of m-open lambda-terms

Theorem. As $n \to \infty$, the number of plain lambda-terms of size n is asymptotically

$$\frac{b_\infty n^{-3/2}}{2\sqrt{\pi}} \left(\frac{1}{\rho}\right)^n, \quad (1-\rho)^3 = 4\rho^2 \ .$$

Theorem. The asymptotic probability that a random plain lambda-term of size n is m-open tends to some positive constant p_m as $n \to \infty$. This distribution is *computable.*

Asymptotic number of m-open lambda-terms

Theorem. The asymptotic probability that a random plain lambda-term of size n is m-open tends to some positive constant p_m as $n \to \infty$. This distribution is *computable*.

Open question. What is the behaviour of the sequence $(p_k)_{k=0}^{\infty}$ which is a cumulative distribution function? We only know that $p_{m+1} \geq p_m$ and $p_m \to 1$ as $m \to \infty$.

Asymptotic number of m-open lambda-terms

Theorem. The asymptotic probability that a random plain lambda-term of size n is m-open tends to some positive constant p_m as $n \to \infty$. This distribution is *computable*.

Open question. What is the behaviour of the sequence $(p_k)_{k=0}^{\infty}$ which is a cumulative distribution function? We only know that $p_{m+1} \geq p_m$ and $p_m \to 1$ as $m \to \infty$. The reccurence can be considered either forward or backwards:

$$\begin{cases} a_{m+1} &= a_m/\rho - a_m^2 - \dfrac{1-\rho^m}{1-\rho}, \\ b_{m+1} &= b_m/\rho - 2a_m b_m, \\ p_m &= b_m/b_\infty \end{cases}$$

Outline

Zoo of different statistics

Marking techniques
Number of lambdas
Number of variables
Number of abstractions
Number of redexes
Value of de-Bruijn index
Number of head abstractions

Extremal techniques
Maximal de-Bruijn index value
Unary height of a random term
Longest lambda-run

Advanced marking
Number of free variables
Number of closed subterms
Expected search time for β-reduction
Number of variables bound to top lambda

Marking techniques

Marking techniques
Number of lambdas
Number of variables
Number of abstractions
Number of redexes
Number of head abstractions
Value of de-Bruijn index

Statistics with gaussian distribution

$\boxed{\textbf{Theorem.}}$ In random plain lambda-term of size n

- $\mathbb{E}_n(\#\ \text{lambdas}) = C_\lambda \cdot n + O(n^{1/2})$,
- $\mathbb{E}_n(\#\ \text{applications}) = C_@ \cdot n + O(n^{1/2})$,
- $\mathbb{E}_n(\#\ \text{variables}) = C_\mathcal{N} \cdot n + O(n^{1/2})$,
- $\mathbb{E}_n(\#\ \text{redexes}) = C_{\text{redex}} \cdot n + O(n^{1/2})$,

The distribution is asymptotically Gaussian, i.e.

$$\boxed{\frac{\# - \mathbb{E}\#}{\mathbb{V}\#} \xrightarrow{d} \mathcal{N}(0,1)}$$

Statistics with gaussian distribution

$\boxed{\textbf{Theorem.}}$ In random **closed** lambda-term of size n

- $\mathbb{E}_n(\#\text{ lambdas}) = C_\lambda \cdot n + O(n^{1/2})$,
- $\mathbb{E}_n(\#\text{ applications}) = C_@ \cdot n + O(n^{1/2})$,
- $\mathbb{E}_n(\#\text{ variables}) = C_\mathcal{N} \cdot n + O(n^{1/2})$,
- $\mathbb{E}_n(\#\text{ redexes}) = C_{\text{redex}} \cdot n + O(n^{1/2})$,

The constants C_λ, $C_@$, $C_\mathcal{N}$, C_{redex} are the same as in the plain case.

$$\boxed{\frac{\# - \mathbb{E}\#}{\mathbb{V}\#} \xrightarrow{d} \mathcal{N}(0,1)}$$

Head abstractions

Theorem.

- The number of head abstractions in random plain lambda-term has a limiting distribution $Geom(\rho)$, i.e.

$$\mathbb{P}(\#\text{ head abstractions } \leq m) = 1 - \rho^{m+1}$$

- The number of head abstractions in random **closed** lambda-term has cumulative distribution function

$$\mathbb{P}(\#\text{ head abstractions } \leq m) = 1 - \rho^{m+1}\frac{p_{m+1}}{p_0}$$

Head abstractions

Theorem.

- The number of head abstractions in random plain lambda-term has a limiting distribution $Geom(\rho)$, i.e.

$$\mathbb{P}(\#\text{ head abstractions } \leq m) = 1 - \rho^{m+1}$$

- The number of head abstractions in random **closed** lambda-term has cumulative distribution function

$$\mathbb{P}(\#\text{ head abstractions } \leq m) = 1 - \rho^{m+1}\frac{p_{m+1}}{p_0}$$

Distribution of de-Bruijn index

Theorem.

- In large random plain lambda-terms, the value of de Bruijn index has a limiting distribution which is *Geom*(ρ).
- In large random **closed** lambda-terms, the value of de Bruijn index has a *computable* limiting distribution.

$$\mathbb{P}_m \sim \frac{\rho^{2m}}{(1 - 2\rho a_0) \dots (1 - 2\rho a_{m-1})}$$

Distribution of de-Bruijn index

Theorem.

- In large random plain lambda-terms, the value of de Bruijn index has a limiting distribution which is $Geom(\rho)$.
- In large random **closed** lambda-terms, the value of de Bruijn index has a *computable* limiting distribution.

$$\mathbb{P}_m \sim \frac{\rho^{2m}}{(1-2\rho a_0)\dots(1-2\rho a_{m-1})}$$

Advanced marking

Advanced marking
Number of free variables
Number of closed subterms
Expected search time for β-reduction
Number of variables bound to head lambda

Number of free variables

Theorem. Inside large plain lambda-terms the number of free variables has a *computable* discrete limiting distribution, in particular, the average number of free variables is a constant

$$\mathbb{E}_n \sim \frac{2}{(1-\rho)^3}$$

Number of free variables

Theorem. Inside large plain lambda-terms the number of free variables has a *computable* discrete limiting distribution, in particular, the average number of free variables is a constant

"Paradox" Almost all the variables are bounded but not all the terms are closed!

Number of free variables

Theorem. Inside large plain lambda-terms the number of free variables has a *computable* discrete limiting distribution, in particular, the average number of free variables is a constant

"Paradox" Almost all the variables are bounded but not all the terms are closed!

Intuition. Distribution of db-index is geometric, but unary height is $O(\sqrt{n})$. A small proportion of variables makes the term open with positive probability.

Number of closed subterms

Theorem. Inside large plain (also closed) lambda-terms the number of closed subterms satisfies

$$\mathbb{E}_n \sim \Theta(n), \quad \mathbb{V}_n \sim \Theta(1)$$

Open question. What is the distribution?

Expected search time for β-reduction

Theorem. Inside large plain lambda-terms the number of steps until first redex discovery has a *computable* discrete limiting distribution, in particular, the expected time is a computable constant.

Average number of variables bound to top lambda

Theorem. Inside large plain lambda-terms choose an abstraction uniformly at random among abstractions at unary height 1.

- $\mathbb{E}$ (number of vars bound by top lambda) $\sim C$
- The same holds for lambda at fixed unary height m. The constant is not necessary the same.

Open question 1.

- Limiting discrete distribution?
- The same holds for closed lambda-terms?

Open question 2.* Number of binding lambdas?

Extremal techniques

- Maximal de-Bruijn index value
- Unary height of a random term
- Longest lambda-run

Extremal statistics

Theorem. In random plain lambda-term of size n

- $\mathbb{E}_n(\text{longest lambda-run}) \sim \dfrac{\log n}{\log(1/\rho)} + O(\log\log n)$,
- $\mathbb{E}_n(\text{maximal de Bruijn index}) \sim \dfrac{\log n}{\log(1/\rho)} + O(\log\log n)$,
- $\mathbb{E}_n(\text{unary height}) \sim \Theta(\sqrt{n})$.

Conjecture. The same is true for closed lambda-terms.

Outline

Open problems

Supposedly hard

- Combinatorics of lambda-term after beta-reduction procedure
- Closed BCI, BCK, λ-I terms.
  Each lambda binds $(\cdot)$ variables
  - ≤ 1 BCI
  - $= 1$ λ-I
  - ≥ 1 BCK

  Riccati PDE, multivariate saddle-point, etc.

  [Lescanne '17] SwissCheese: keeping a large vector of information to track the number of variables on every level
- Number of binding lambdas
- Phase transitions with respect to *abcd* size notion

Open problems

Supposedly hard

- Combinatorics of lambda-term after beta-reduction procedure
- Closed BCI, BCK, λ-I terms.
  Each lambda binds $(\cdot)$ variables
  - ≤ 1 BCI
  - $= 1$ λ-I
  - ≥ 1 BCK

  Riccati PDE, multivariate saddle-point, etc.

  [Lescanne '17] SwissCheese: keeping a large vector of information to track the number of variables on every level
- Number of binding lambdas
- Phase transitions with respect to *abcd* size notion

Open problems

Supposedly hard

- Combinatorics of lambda-term after beta-reduction procedure
- Closed BCI, BCK, λ-I terms.
  Each lambda binds $(\cdot)$ variables
  - ≤ 1 BCI
  - $= 1$ λ-I
  - ≥ 1 BCK

  Riccati PDE, multivariate saddle-point, etc.

  [Lescanne '17] SwissCheese: keeping a large vector of information to track the number of variables on every level
- Number of binding lambdas
- Phase transitions with respect to *abcd* size notion

Open problems

Supposedly hard

- Combinatorics of lambda-term after beta-reduction procedure
- Closed BCI, BCK, λ-I terms.
  Each lambda binds $(\cdot)$ variables
  - ≤ 1 BCI
  - $= 1$ λ-I
  - ≥ 1 BCK

  Riccati PDE, multivariate saddle-point, etc.

  [Lescanne '17] SwissCheese: keeping a large vector of information to track the number of variables on every level
- Number of binding lambdas
- Phase transitions with respect to *abcd* size notion

That's all!

Thank you for your attention!